Title

The argument-adjunct asymmetry revisited: The role of focus alternatives in island effects

Permalink

https://escholarship.org/uc/item/84k014xv

Authors

Yue, Jiayuan

Xiang, Ming

Publication Date

2025

DOI

10.7280/S9XD0ZQP

Peer reviewed

The argument-adjunct asymmetry revisited:
The role of focus alternatives in island effects

Jiayuan Yue & Ming Xiang[*]

Abstract. The unacceptability of filler-gap dependencies in island constructions has been attributed to multiple factors including syntactic constraints, processing difficulty, and discourse conditions. This study examines Chinese *wh*-questions in relative clause islands (RC-islands). While Chinese adjunct *wh*-questions have been shown to be sensitive to RC-islands, whether argument *wh*-questions are is still under debate. We explore the hypothesis that the unacceptability of *wh*-questions with RC-islands arises from the difficulty of generating a set of focus alternatives relevant for the question. In a sentence acceptability experiment, we manipulated the availability of context – specifically the availability of focus alternatives that may serve as answers to the target *wh*-question. The focus alternatives reduced the sensitivity of argument *wh*-questions to RC-islands, but not the adjunct *wh*-questions. We discuss the implications of such contextual effects on the discourse approach to islands.

Keywords. island effect; focus; *wh*-questions; discourse structure; superadditivity; Chinese

1. Introduction. In a filler-gap construction, a constituent (the "filler") is displaced from its canonical position (the "gap"). A common example is *wh*-questions, such as "what did William buy ___?", where the *wh*-word appears at a frontal position instead of its canonical postverbal position indicated by the underscore. Other examples include relative clauses (RC), topicalization, it-clefts, etc. It has long been noticed that some filler-gap dependencies are unacceptable (Ross 1967; Chomsky 1973, among others). In (1), the extraction of the *wh*-element from a relative clause is unacceptable. Many other syntactic environments also render a *wh*-dependency unacceptable, including complex NPs, adjuncts, coordination, etc. These syntactic environments are called "islands" – extractions from such islands yield ill-formed filler-gap dependencies.

(1)   *What did you see the scientist who invented ___?

Explaining the unacceptability of these filler-gap dependencies is a question under debate. Many different accounts have been proposed based on syntactic theories, sentence processing mechanisms, and discourse structure (see Liu et al. 2022 for a review). This paper studies the effects of the discourse context on the acceptability of Chinese *wh*-questions with relative clause islands (RC-islands). Building on previous accounts of island effects based on discourse theories of focus and background, we experimentally manipulated the contextual availability of focus alternatives and collected acceptability ratings of *wh*-questions. We found that making the relevant focus alternatives salient in the context improved the acceptability of argument *wh*-questions with RC-islands, but not adjunct *wh*-questions. We discuss the implications of these results on the discourse approach to island effects.

1.1. SYNTACTIC APPROACHES TO ISLAND: THEORIES AND EXPERIMENTS. One prominent approach to explain the island phenomenon with formal syntax is the subjacency condition (Chomsky 1973; Huang 1982; Rizzi 1982, among others). In a *wh*-question, the syntactic operation of movement extracts the *wh*-word from a lower position to a higher position in the left periphery. This movement may be constrained by the intervening nodes between the two positions. Chomsky (1973) identified S and NP as two types of bounding nodes that restrict syntactic movements. If there are two or more bounding nodes intervening between the original position and the target position, the movement is ungrammatical. In this way, the grammar prevents the extraction of *wh*-elements from island constructions. For instance, the *wh*-extraction in (1), repeated in (2), is ungrammatical since there are three intervening bounding nodes on the path of the movement.

(2)   *What [$_S$ did you meet [$_{NP}$ the scientist who [$_S$ invented ___?]]]

However, many factors may contribute to the unacceptability of *wh*-questions like (1) besides grammatical constraints. For instance, the sheer presence of the relative clause structure could make the sentence less acceptable, as relative clauses are difficult to parse and require memory encoding, maintenance, and retrieval of the filler. The length of the movement in (1) could be another factor. Longer dependencies were shown to be more costly for working memory processing (Phillips et al. 2005). The relative clause structure and dependency length together might be sufficient to explain the unacceptability of (1), without the need for any further syntactic constraints.

Experimental works were conducted to eliminate the two confounds – structure and dependency length – and to establish a quantitative factorial definition of islands (Sprouse et al. 2012, 2016). Sprouse et al. (2016) conducted a series of sentence acceptability ratings tasks in English and Italian. They tested the acceptability of *wh*-dependencies and RC-dependencies in various island constructions. For English complex NP islands, their experiment had a 2x2 design, with two levels for Structure (embedded clause vs. complex NP) and two levels for Dependency Length (short dependency vs. long dependency). An example item is shown in (3). We note that although (3b) and (3d) in this design are both classified as "Long", the dependency in (3d) is even longer than the dependency (3b). The same design can be easily adopted for other dependencies, island types, and languages. The key prediction was: if island effects arise from special syntactic constraints like the Subjacency Condition, then the unacceptability of (3d) should not just be an additive sum of the main effects of Structure and Dependency length, but should show a superadditive interaction between the two factors.

(3)   (Sprouse et al. 2016: 318)
       a.   Short, Embedded: Who ___ heard that Jeff baked a pie?
       b.   Long, Embedded: What did you hear that Jeff baked ___?
       c.   Short, Complex NP: Who ___ heard the statement that Jeff baked a pie?
       d.   Long, Complex NP: What did you hear the statement that Jeff baked ___?

The authors found a significant interaction between Structure and Dependency length that particularly impacted the ratings of the Long Dependency, Complex Structure condition (3d). The interaction shows that the complex NP structure and the dependency length alone were not sufficient to explain the island effect, calling for additional explanations such as the subjacency condition. This interaction was found in most of the island structures identified in previous syntactic literature, but was absent in some others, including RC-dependencies in adjunct islands in English, such as "I called the client who the secretary worries if the lawyer insults __" (Sprouse

et al. 2016). The authors interpreted the absence of superadditivity as evidence that English relative clauses do not show true adjunct island effects.

1.2. NON-SYNTACTIC APPROACHES TO ISLAND: PROCESSING AND DISCOURSE. While superadditivity shows there are additional factors behind island effects besides complex structure and dependency length, the additional factors are not necessarily syntactic. The sentence processing approach may attribute the unacceptability of sentences like (3d) to the particularly high cost of memory retrieval (Kluender & Kutas 1993; Keshev & Meltzer-Asscher 2019). When processing a *wh*-question, the intervening NPs between the *wh*-word and the gap could trigger memory interference at the retrieval site. In the short-distance dependency conditions (3a) and (3c), the *wh*-word was right next to the gap with no intervening NPs. The Long, Embedded Clause condition (3b) had two intervening NPs – the matrix subject NP and the embedded subject NP. The Long, Complex NP condition (3d) had three intervening NPs – the matrix subject NP, the matrix object NP, and the relative clause subject NP. This additional intervening NP could be the source of the Structure x Length interaction. These intervening NPs interfere in memory retrieval as we search for the correct retrieval target. The retrieval difficulty would be the highest in a superadditive way for the Long, Complex NP condition.

Another approach to explaining the island effects focuses on the discourse function of the sentence in its context. Early theories posit discourse constraints on the constituents that are allowed to move, including the Dominance condition and the Topichood condition (Erteschik-Shir 1973, 1986; Kuno 1987). A constituent is dominant if the speaker wishes to draw attention to it (Erteschik-Shir 1986). This notion of dominance is somewhat related to the concept of discourse focus, which is the new information provided by the speakers in an utterance. The Dominance condition on movement proposes that "an NP can only be extracted out of clauses which may be interpreted as dominant, or out of phrases in which the NP may itself be dominant" (Erteschik-Shir 1986). Therefore, if an island structure is unlikely to be what the speaker is drawing attention to, then the Dominance condition could explain the unacceptability of wh-movements out of such structures. In a similar vein, the Topichood condition restricts that "only those constituents in a sentence that qualifies as the topic of the sentence can undergo extraction processes" (Kuno 1987). A topic is an NP that is likely to be the subject of the next sentence, thus reflecting what the conversation is about. Synthesizing these two early accounts, Goldberg (2013) argues that both focus constituents and topic constituents are discourse-prominent, as opposed to the rest of the sentence that is backgrounded. The Backgrounded Constituents are Islands (BCI) constraint proposes that backgrounded constituents cannot be extracted in long-distance dependencies. Some recent work also proposed the Focus-Background Conflict (FBC) (Abeillé et al. 2020), which states that "a focused element should not be part of a backgrounded constituent". For example, the subject of a sentence is typically a backgrounded constituent, as it is often part of the given information based on which new comments will follow. Therefore, *wh*-questions should be sensitive to the subject island, as there is a conflict between the focused *wh*-word and the backgrounded subject constituent. Relative clauses, without focus on the head NP, should not be sensitive to subject island. Results from acceptability rating experiments in Abeillé et al. (2020) indeed support this account.

2. Island effect in Mandarin Chinese. Mandarin Chinese is a *wh*-in-situ language. Syntactic work on *wh*-questions in Chinese has generally argued that argument *wh*-questions in Chinese are not sensitive to island constraints, while adjunct *wh*-questions are (Huang 1982; Tsai 1994, among others). Typical argument *wh*-words in Chinese include *shenme* 'what' and *shei* 'who',

and typical adjunct *wh*-words in Chinese include *weishenme* 'why' and *zenme* 'how'. In (4) we show some examples of *wh*-questions with relative clauses. The contrast between the acceptable argument *wh*-question in (4a) and the unacceptable adjunct *wh*-question in (4b) shows the asymmetry between the two *wh*-categories.

(4)   a.   玛丽 见 了   研究 什么   的 人
            Mali jian le [NP yanjiu shenme de ren]?
            Mali see ASP   study   what   REL person
            'What did Mali see a person that studies ___?'
    b.   玛丽 见 了   为什么   研究 数学   的   人
        *Mali jian le [NP weishenme yanjiu shuxue de   ren]?
            Mali see ASP   why   study math   REL person
            'Why did Mali see a person that studies math ___?'
            Intended answer: They study math because they love geometry.

Since Chinese is a *wh*-in-situ language where the *wh*-word does not overtly move out of its canonical position, syntactic approaches to Chinese islands focus on covert movements. The presence or absence of covert movement could be a possible explanation for the argument-adjunct asymmetry. Tsai (1999) argues that *wh*-arguments do not go through covert phrasal movement. Instead, the question operator binds to the *wh*-element through unselective binding. Therefore, they are not sensitive to the Subjacency Condition on movements. In example (4a), this Q operator would be at the left edge of the matrix clause, which determines the semantic scope for the in-situ *wh*-phrase. Unselective binding is not available for *wh*-adjuncts, which have to undergo covert phrasal movement, resulting in island sensitivity. An alternative theory of *wh*-in-situ, without invoking covert movement, employs choice functions to account for the argument-adjunct asymmetry in English multiple-*wh* questions, where one *wh*-phrase stays in-situ (Reinhart 1998). A key distinction was made between *wh*-arguments and *wh*-adjuncts such that only *wh*-arguments can be interpreted via choice functions.

    Although *wh*-words in Chinese *wh*-questions remain in-situ, comprehension of such questions is not immune to the aforementioned processing costs in working memory retrieval (Xiang et al. 2014, 2015). Under the assumption mentioned above that the in-situ *wh*-element is associated with a silent Q-operator at its semantic scope position, the processing speed of the *wh*-word was affected by the length of the covert dependency and the number of intervening items. Therefore, the acceptability of Chinese *wh*-questions, like their English and Italian counterparts, may also be sensitive to the processing difficulty of the complex structure and the long-distance dependency, independent of any further syntactic constraints. Superadditivity, as a measurement of the island effect, is necessary to rule out these processing cost factors.

    Experimental tests of Chinese islands, adopting Sprouse et al.'s (2016) superadditivity design, have yielded mixed results (Lu et al. 2020; Tian et al. 2022). Lu et al. (2020) conducted an acceptability judgment experiment on Chinese *wh*-questions with RC-islands in a 2x2x2 design. Besides the 2x2 design of superadditivity (Structure x Length), they added the additional variable of Wh-category, testing both argument *wh*-questions (*shenme* 'what', *shei* 'who') and adjunct *wh*-questions (*weishenme* 'why'). Tables 1 and 2 show an example item used in the study. If there were an argument-adjunct asymmetry in the sensitivity of Chinese *wh*-questions to RC-islands, then adjunct *wh*-questions should show a superadditivity interaction, while argument *wh*-questions should not. On the contrary, their results showed superadditivity for both argument and

adjunct *wh*-questions, with no significant three-way interaction between Structure, Length, and Wh-category. In other words, *wh*-category did not affect *wh*-questions' sensitivity to RC-islands.

Condition	Sentence								
Short, Embed	约翰	想知道	谁	说	女孩	吃	了	寿	
	Yuehan	xiangzhidao	shei	shuo	nyuhai	chi	le	shousi	
	John	wonders	who	say	girl	eat	ASP	sushi	
	'John wonders who ___ said that the girl ate sushi.'								
Short, RC	约翰	想知道	谁	见	了	吃	寿司	的	女孩
	Yuehan	xiangzhidao	shei	jian	le	chi	shousi	de	nyuhai
	John	wonders	who	meet	ASP	eat	sushi	REL	girl
	'John wonders who ___ met the girl that ate sushi.'								
Long, Embed	约翰	想知道	比尔	说	女孩	吃	了	什么	
	Yuehan	xiangzhidao	Bier	shuo	nyuhai	chi	le	shenme	
	John	wonders	Bill	say	girl	eat	ASP	what	
	'John wonders what Bill said that the girl ate ___.'								
Long, RC	约翰	想知道	比尔	见	了	吃	什么	的	女孩
	Yuehan	xiangzhidao	Bier	jian	le	chi	shenme	de	nyuhai
	John	wonders	Bill	meet	ASP	eat	what	REL	girl
	'John wonders what Bill met the girl that ate ___.'								

Table 1. Example stimuli for argument *wh*-questions in Lu et al. (2020)

Condition	Sentence							
Short, Embed	约翰	想知道	比尔	为什么	说	女孩	吃 了	寿司
	Yuehan	xiangzhidao	Bier	weishenme	shuo	nyuhai	chi le	shousi
	John	wonders	Bill	why	say	girl	eat ASP	sushi
	'John wonders why Bill says ___ that the girl ate sushi.'							
Short, RC	约翰	想知道	比尔	为什么	见 了 吃	寿司	的	女孩
	Yuehan	xiangzhidao	Bier	weishenme	jian le chi	shousi	de	nyuhai
	John	wonders	Bill	why	meet ASP eat	sushi	REL	girl
	'John wonders why Bill met ___ the girl that ate sushi.'							
Long, Embed	约翰	想知道	比尔	说	女孩	为什么	吃 了	寿司
	Yuehan	xiangzhidao	Bier	shuo	nyuhai	weishenme	chi le	shousi
	John	wonders	Bill	say g	irl	why	eat ASP	sushi
	'John wonders why Bill says that the girl ate sushi ___.'							
Long, RC	约翰	想知道	比尔	见	了 为什么	吃 寿司	的	女孩
	Yuehan	xiangzhidao	Bier	jian	le weishenme	chi shousi	de	nyuhai
	John	wonders	Bill	meet	ASP why	eat sushi	REL	girl
	'John wonders why Bill met the girl that ate sushi ___.'							

Table 2. Example stimuli for adjunct *wh*-questions in Lu et al. (2020)

A potential issue with the materials used by Lu et al. (2020) was pointed out by Tian et al. (2022). Tian et al. (2022) argue that relative clauses tend to "characterize the prominent feature of the relativized nominal head". For example, a discourse-prominent feature of hunters may be the kind of prey that they hunt. A relative clause such as "the hunter who hunts deer" is pragmatically felicitous. By contrast, eating something is usually not a prominent feature of a girl, thus

the relative clause "the girl who eats sushi" is infelicitous in an out-of-the-blue context. In this way, pragmatic felicity may be a factor driving down the acceptability of *wh*-questions in Lu et al.'s (2020) study. Tian et al. (2022) replicated this study with new materials. They constructed *wh*-questions with relative clauses that characterize discourse-prominent features, such as those in (5). Other examples include "the news that reports on …", "the policeman that inspects …", and "the salesperson that is good at selling …". The results showed a superadditivity interaction for adjunct *wh*-questions, but not argument *wh*-questions. There was a three-way interaction between Structure, Length, and Wh-category, showing that Wh-category indeed affected questions' sensitivity to islands. Therefore, the results of Tian et al. (2022) support an argument-adjunct asymmetry, contrary to Lu et al. (2020). The contrast between the results of the two studies points to the importance of "discourse prominence" in the acceptability of *wh*-questions in Chinese.

(5) 周勇   想知道   政府   会   惩罚   捕杀   什么   的 猎人

Zhouyong xiangzhidao zhengfu hui chengfa busha shenme de lieren

Zhouyong wonder government will punish hunt what REL hunter

Intended: 'Zhouyong wonders which kind of hunters the government will punish.'

3. Hypothesis: Availability of focus alternatives. The notion of "discourse prominence", however, is not clearly defined in Tian et al. (2022). One way to formulate this concept more clearly is to define it as the availability of alternatives. The discourse-prominent relative clauses constructed by Tian et al. (2022) have easily accessible alternatives. For example, a relative clause like "a teacher that teaches math" could easily be substituted by "a teacher that teaches English", "a teacher that teaches physics", or "a teacher that teaches history". All these alternatives are highly accessible to the interlocutors because when we think of teachers, since we often think of what subject they teach. Similarly, "a poacher that hunts whales" can easily be replaced by "a poacher that hunts rhinos", or "a poacher that hunts elephants". On the other hand, it's less likely that when we think of a girl, we think of different food options that she eats. Therefore, alternatives in the materials used by Lu et al. (2020) are less available to participants in the out-of-the-blue context. In this way, the diverging results on the argument-adjunct asymmetry may be attributed to the accessibility of alternatives.

This informal idea of the availability of alternatives can be further developed using the formal model of focus alternatives by Rooth (1985, 1992). While the sentence "[Mary]$_F$ likes Sue" with the focus on "Mary" and a different sentence "Mary likes [Sue]$_F$" with the focus on "Sue" have the same ordinary semantic value in terms of truth conditions, they have different focus semantic values. The focus semantic value of a sentence is a set of propositions, where the focus position is replaced by possible alternatives, and the rest of the proposition is kept the same. Therefore, as shown in (6), the focus semantic value of "[Mary]$_F$ likes Sue" is a set of propositions in the form of "X likes Sue", where X can be replaced by possible individuals. The focus semantic value of "Mary likes [Sue]$_F$" is a set of propositions in the form of "Mary likes Y", where Y can be replaced by possible individuals.

(6) a. $[[Mary]_F \text{ likes Sue}]^f = \{like(x, Sue) \mid x \text{ in } E\}$ where E is the domain of individuals
     b. $[Mary \text{ likes } [Sue]_F]^f = \{like(Mary, y) \mid y \text{ in } E\}$

This concept of focus as alternatives is particularly useful for our current purpose. The *wh*-word in *wh*-questions is a focused element. It initiates the construction of a set of possible alternatives based on the discourse context, and then allows for evaluations of each element of the set in

relation to the rest of the sentence. For example, the question "who won the race?" has its focus on "who". There is a set of possible discourse referents, such as {John, Mary, William}. To answer this *wh*-question, one would need to generate the alternatives set {John won the race, Mary won the race, William won the race}. If John won the race is true, then the interlocutor can answer "John won the race" with the focus on "John" as a selection from the alternative set. Therefore, the construction of the alternative set is crucial to the interpretation of *wh*-questions.

The Focus-Background Conflict account of island effects can also be further elaborated using focus alternatives. Since a focused element is unexpected to appear in the backgrounded part of a sentence, it is harder to construct the proper focus alternative set for a sentence that violates the FBC. Therefore, *wh*-questions with islands violate the reader's expectations. Readers could be confused about the information structure of the sentence. It would be difficult for a reader to interpret what kind of new information the question is inquiring about, and what part of the sentence is presupposed or given. This interpretation difficulty may affect the acceptability ratings of *wh*-questions. If the focus-background partition of a *wh*-question is clear and the reader can easily construct a set of relevant alternatives, the acceptability of the *wh*-question should improve.

To summarize, we hypothesize that the comprehension of *wh*-questions requires the construction of a set of contextually salient alternatives. Easily accessible alternatives may facilitate the interpretation of *wh*-questions and reduce island effects. The availability of focus alternatives can be manipulated experimentally by controlling the context of the sentence. While it is difficult to think of alternatives for "a girl that eats sushi" in an out-of-the-blue context, one can construct a context that mentions a set of girls that are eating different food, such as "a girl that eats pizza" and "a girl that eats salads". Having these accessible alternatives in the context may help the reader construct focus alternatives and therefore facilitate the focus-background partition in the interpretation process of the *wh*-question. Therefore, providing a set of alternatives in context may improve the acceptability of *wh*-extraction out of relative clauses in Chinese. We test this for both argument and adjunct *wh*-questions.

4. Acceptability rating experiment. We conducted an acceptability judgment experiment on Mandarin *wh*-questions, including both argument *wh*-questions and adjunct *wh*-questions. This experiment aims to study the effect of discourse context, specifically the availability of focus alternatives, on the acceptability of Chinese *wh*-questions with RC-islands.

4.1. DESIGN. This study replicates Lu et al.'s (2020) design of acceptability judgments and adds one additional variable: Context. Taken together, the experiment has a 2x2x2x2 design (Context x Structure x Length x Wh-category). Context is a between-subject variable with two levels: No Context vs. With Context. In the With Context condition, the participant reads a context description before rating the target sentence. The specifics of the provided context are described in the Materials section. In the No Context condition, the participant simply rates target sentences. Structure is a within-subject variable with two levels: Embedded clause vs. Relative clause. Length of dependency is a within-subject variable with two levels: Short vs. Long. Finally, Wh-category is a within-subject variable with two levels: Wh-argument (what, who) and Wh-adjunct (why). Structure and Length together help us examine the superadditivity effect of islands introduced in section 1.

4.2. HYPOTHESIS. We hypothesize that the sensitivity of Chinese *wh*-questions to RC-islands is due to difficulties in the informational structure partition of focus and background. Having a set of focus alternatives available in the discourse context that could potentially serve as answers to

the *wh*-question would facilitate an easier partition of its information structure. Therefore, increased availability of focus alternatives should improve the acceptability of Chinese *wh*-extraction out of islands.

Regarding the argument-adjunct asymmetry, we hypothesize, following Lu et al. (2020), that both categories of *wh*-questions are sensitive to RC-islands in the absence of a supporting context. Two competing hypotheses can be made on the effect of discourse context on argument *wh*-questions vs. adjunct *wh*-questions. On the one hand, since both *wh*-arguments and *wh*-adjuncts are interpreted by alternative sets, it is possible that both type of *wh*-questions are equally affected by the availability of alternatives in the discourse context. On the other hand, since the alternatives for *wh*-arguments are different from the alternatives for *wh*-adjunct, it is also possible that the discourse context has stronger effects on one *wh*-category than the other.

4.3. MATERIALS. The target sentences (24 sets) for the acceptability judgment were adapted and modified from Lu et al. (2020). The major modification in the material is that we turned the target *wh*-dependency into a matrix question instead of keeping it as an embedded question. For example, a target sentence like "John wonders who said the boys bought toys" was changed to "who said the boys bought toys?" In this way, we keep the materials in accordance with Sprouse et al. (2016) and Abeillé et al. (2020). This change generally reduces the processing difficulty of the materials. And an explicit matrix question also makes the focus status of the *wh*-element more salient.

In the With-Context condition, the contexts before the target sentences were designed to provide two alternatives for the *wh*-question. For example, if the *wh*-question was "who said that the boys bought toys?", the context would mention two individuals, such as a cashier who said the boys bought toys and a manager who said the boys bought comic books. Then the reader would have access to the alternatives set {the cashier said that the boys bought toys, the manager said that the boys bought toys} when they tried to interpret the *wh*-question. The contexts started with a general topic sentence, such as "some boys are shopping in a grocery store." Then the alternatives were given after this topic sentence.

The lexical items and sentence structure of the context sentences were carefully chosen to minimize overlap with the target sentence. For example, if the target sentence was "who said the boys bought toys?", then we avoided using the same verb "bought" in the context and replaced it with verbs like "picked up". With this manipulation, any improvement in the acceptability of the target sentence could not be explained by the lexical repetition of the same words and syntactic structures.

Examples of experiment items are shown in Tables 3 and 4, including both the context and the target sentence. The context is only shown to the participants in the With Context condition. The target sentence is shown to all participants.

4.4. PARTICIPANTS. Fifty-seven Chinese native speakers were recruited from social media and participated in this experiment (23 males; mean age = 21.8). The experiment was advertised on social media groups for Chinese students in various American universities. Each of them was compensated $5 in the form of Amazon gift cards after completion of the study. Twenty-six participants were randomly assigned to the No Context condition, and 31 participants were assigned to the With Context condition.

4.5. PROCEDURE. For the between-participant variable Context (context vs. no context), participants were randomly assigned to one of the two experiment groups on PCIbex (Zehr & Schwarz 2018), either the No Context group or the With Context group. Under each group, the 24 sets of

experimental items were distributed in a Latin-square manner based on the 2x2x2 design, with each participant completing 3 trials per condition (24 experimental trials total). The experiment took approximately 20 minutes.

Condition	Sentence
Short, Embed	几个工人在客厅里修家具。许厂长表示工人们更换了坏掉的窗户。梁师傅说他们把桌子加固了。 'Several workers are repairing furniture in the living room. Director Xu mentioned that the workers replaced the broken window. Craftsman Liang said they strengthened the table.' **Target:** 谁　　说　　工人　　修　了　　桌子？ Shei　shuo　gongren　xiu　le　zhuozi Who　say　worker　repair　ASP　table 'Who ___ said the workers repaired the table?'
Short, RC	几个工人在客厅里修家具。赵徒弟更换了坏掉的窗户，而李徒弟加固了桌子。许厂长称赞了赵徒弟。梁师傅赞美了李徒弟。 'Several workers are repairing furniture in the living room. Apprentice Zhao replaced the broken window, and Apprentice Li strengthened the table. Director Xu complimented Apprentice Zhao. Craftsman Liang praised Apprentice Li.' **Target:** 谁　　赞美　　了　　修　了　　桌子　的　　工人？ Shei　zanmei　le　xiu　le　zhuozi　de　gongren Who　praise　ASP　repair　ASP　table　REL　worker 'Who ___ praised the worker that repaired the table?'
Long, Embed	几个工人在客厅里修家具。许厂长表示工人们更换了坏掉的窗户。梁师傅说他们把桌子加固了。 'Several workers are repairing furniture in the living room. Director Xu mentioned that the workers replaced the broken window. Craftsman Liang said they strengthened the table.' **Target:** 梁　　师傅　　说　　工人　　修　　了　什么？ Liang　shifu　shuo　gongren　xiu　le　shenme Liang　craftsman　say　worker　repair　ASP　what 'What did Craftsman Liang say the workers repaired ___?'
Long, RC	几个工人在客厅里修家具。赵徒弟更换了坏掉的窗户，而李徒弟加固了桌子。许厂长称赞了赵徒弟。梁师傅赞美了李徒弟。 'Several workers are repairing furniture in the living room. Apprentice Zhao replaced the broken window, and Apprentice Li strengthened the table. Director Xu complimented Apprentice Zhao. Craftsman Liang praised Apprentice Li.' **Target:** 梁　　师傅　　赞美　　了　修　了　什么　　的　　工人？ Liang　shifu　zanmei　le　xiu　le　shenme　de　gongren Liang　craftsman　praise　ASP　repair　ASP　what　REL　worker 'What did Worker Liang praise the worker that repaired ___?'

Table 3. An example stimuli set for the argument *wh*-questions with context

Condition	Sentence
Short, Embed	几个工人在客厅里修家具。许厂长说他们更换了坏掉的窗户，因为他看到玻璃上的裂痕没了。梁师傅说他们把桌子加固了，因为他发现桌子不晃了。 'Several workers are repairing furniture in the living room. Director Xu mentioned they replaced the broken window, because he saw the crack on the glass was gone. Craftsman Liang said they strengthened the table, because he found the table was no longer wobbly.' **Target:** 梁　师傅　　为什么　说　工人　　修　了　桌子？ Liang　shifu　　weishenme　shuo　gongren　xiu　le　zhuozi Liang　craftsman　why　　say　worker　repair　ASP　table 'Why did Craftsman Liang say __ that the workers repaired the table?'
Short, RC	几个工人在客厅里修桌子。许厂长称赞了他们，因为修好的桌子质量很高。梁师傅也赞美了他们，因为维修的速度很快。 'Several workers are repairing a table in the living room. Director Xu complimented them, because the repaired table was of high quality. Craftsman Liang also praised them, because the speed of the reparation was fast.' **Target:** 梁　师傅　为什么　　称赞　了　修　了　桌子　的　工人？ Liang　shifu　weishenme　chenzan　le　xiu　le　zhuozi　de　gongren Liang　worker　why　　　praise　ASP　repair　ASP　table　REL　worker 'Why did Craftsman Liang praise __ the workers that repaired the table?'
Long, Embed	几个工人在客厅里修桌子。许厂长说是因为房子主人请他们来修。梁师傅则说是因为他们热心帮助屋里的老人。 'Several workers are repairing a table in the living room. Director Xu mentioned it's because the owner of the house arranged for the reparation. Craftsman Liang said it's actually because they are earnestly helping the seniors in the residence.' **Target:** 梁　师傅　　说　工人　　为什么　修　了　桌子？ Liang　shifu　　shuo　gongren　weishenme　xiu　le　zhuozi Liang　craftsman　say　worker　　why　　repair　ASP　table 'Why did Craftsman Liang say the workers repaired the table __?'
Long, RC	几个工人在客厅里修桌子。赵徒弟来修是因为房子主人请他来。李徒弟则是因为热心帮助屋里的老人。梁师傅赞美了李徒弟。 'Several workers are repairing a table in the living room. Apprentice Zhao came because the owner of the house arranged for the reparation, while Apprentice Li came because he was earnest to help the seniors in the residence. Craftsman Liang praised Apprentice Li' **Target:** 梁　师傅　称赞　了　为什么　修　了　桌子　的　工人？ Liang　shifu　chenzan　le　weishenme　xiu　le　zhuozi　de　gongren Liang　worker　praise　ASP　why　　repair　ASP　table　REL　worker 'Why did Craftsman Liang praise the workers that repaired the table ___?'

Table 4. An example stimuli set for the adjunct *wh*-questions with context

In the No Context group, participants were instructed to rate the acceptability of the target sentences on a 1-7 scale. Then they completed four practice trials that were clearly acceptable or unacceptable and received feedback on their choices. In the With Context group, participants were told they would see a context and a target sentence, and they needed to rate the acceptability of the target sentence on a 1-7 scale. They were specifically instructed not to rate the

congruence between the context and the target sentence, but only the acceptability. Then they completed four practice trials and received feedback on their choices, including one target sentence that was clearly acceptable but incongruent with the context, to ensure participants understood the instructions. After completing the instructions and practice trials, each group of participants completed 24 experimental trials and 24 filler trials. All trials were randomized in their presentation order.

4.6. PREDICTIONS. The key prediction, based on the focus-alternative hypothesis, was that the island effect should be weaker in the With-context conditions than in the No-context conditions. When the focus alternatives are provided in the context, the discourse structure of a *wh*-question should be easier to interpret, even when the *wh*-extraction originates within an island structure, leading to higher acceptability of island sentences. Since superadditivity is our behavioral measure for island effect sensitivity, we predict the superadditivity interaction should be weaker in the With Context condition than in the No Context condition. Regarding the comparison between argument and adjunct questions, we predict that both argument and adjunct *wh*-questions should show island sensitivity in the No Context condition, since this is largely a replication of Lu et al. (2020). The more interesting question is whether we would observe any differences in the With-context conditions.

4.7. RESULTS. Ten out of the total 57 participants were excluded from further analysis based on the following criteria. Participants whose mean response time was more than one standard deviation lower than the mean response time in their respective Context conditions were excluded from the analysis, as the quick response time indicated that they had not read the context and the target sentence carefully. Participants whose accuracy of acceptability judgments in filler trials was lower than 75% were also excluded for the same reason. Individual trials with a response time over 30 seconds were discarded.

The results were analyzed by Linear Mixed-Effects Regression models implemented in R using the lme4 package (Bates et al. 2015) and p-values were supplemented by the lmerTest packages (Kuznetsova et al. 2017). Firstly, four linear mixed-effects models were constructed for questions of each Wh-category with or without contexts. The model structure is shown below:

Rating ~ Structure*Length + (1|Item) + (1+Structure*Length|Participant)

Figure 3 shows the results of the acceptability judgment task on argument *wh*-questions with no context provided. There is a significant Structure main effect (Est = -1.04, se = 0.37, $p < 0.01$) such that questions with embedded clauses are generally rated better than questions with relative clauses. There is also a significant Length main effect (Est = 0.76, se = 0.26, $p < 0.01$) such that short dependencies are generally rated higher than long dependencies. Crucially, there is also a significant superadditivity as defined by the Structure x Length interaction ($\chi^2 = 5.19$, $p < 0.05$). The argument *wh*-questions in Mandarin Chinese are sensitive to RC-island structures.
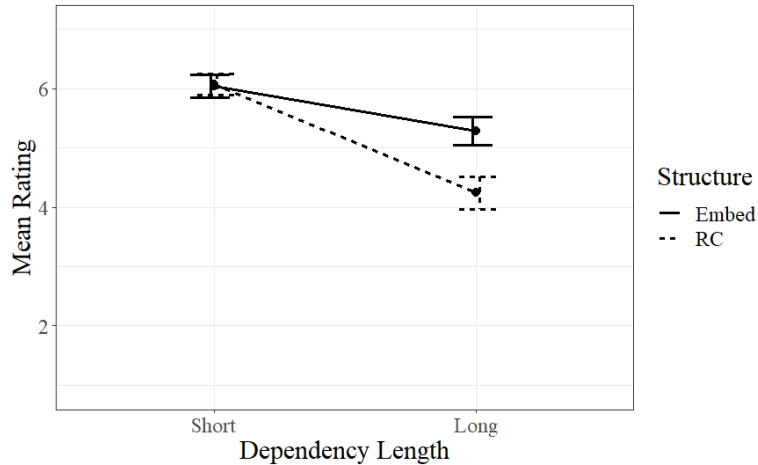
Figure 3. Interaction plot for argument *wh*-questions without context

Figure 4 shows the results of the acceptability judgment task on argument *wh*-questions with a given context that provides a set of focus alternatives. In this condition, there is no significant main effect for either Structure (Est = -0.48, se = 0.43, p > 0.2) or Length (Est = 0.35, se = 0.55, p > 0.5). The Structure x Length interaction is not significant (χ^2 = 3.86, p = 0.05). These results suggest that with the help of context, argument *wh*-questions do not show RC-island sensitivity. In particular, with an appropriate context, the mean acceptability ratings for the island condition (the long RC condition) became very close to its length-matched non-island-extraction counterpart (the long embedded clause condition) (Est = -0.53, se = 0.43, p > 0.2). It is worth noticing that the interaction may appear to be close to significance, but this should not be interpreted as showing an island sensitivity. Instead, the short-dependency questions with embedded vs. RC structures had significantly different ratings (Est = 1.00, se = 0.45, p < 0.05). Therefore, the Structure x Length interaction that is close to significance rather reflects the surprising result that RC structures are more acceptable than embedded structures with a short *wh*-dependency. We do not have any hypothesis for why the two short conditions had different ratings, and we leave it for future work.
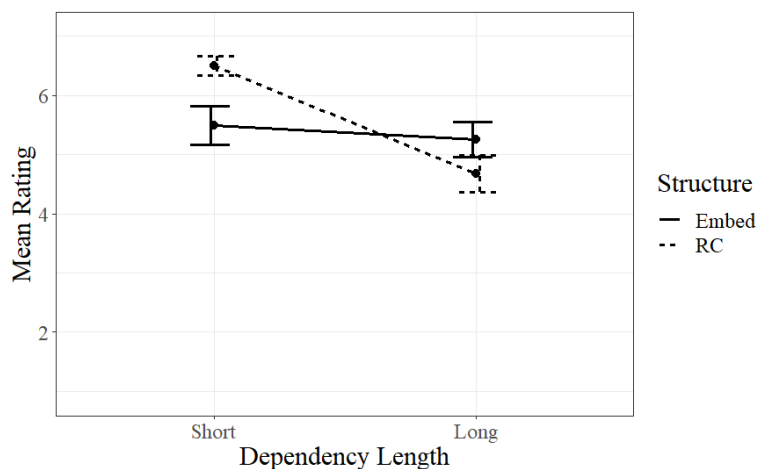

Figure 4. Interaction plot for argument *wh*-questions with context

Figure 5 shows the results of the acceptability judgment task on adjunct *wh*-questions with no context provided. There is a significant Structure main effect (Est = -2.80, se = 0.34, p < 0.001)

578

such that questions with embedded clauses are generally rated better than questions with relative clauses. There is also a significant Length effect (Est = 0.73, se = 0.33, p < 0.05) such that short dependencies are generally rated higher than long dependencies. Crucially, there is also a significant superadditivity as defined by the Structure x Length interaction (χ^2 = 33.71, p < 0.001). When no context is provided, adjunct *wh*-questions in Chinese are strongly sensitive to RC-islands.
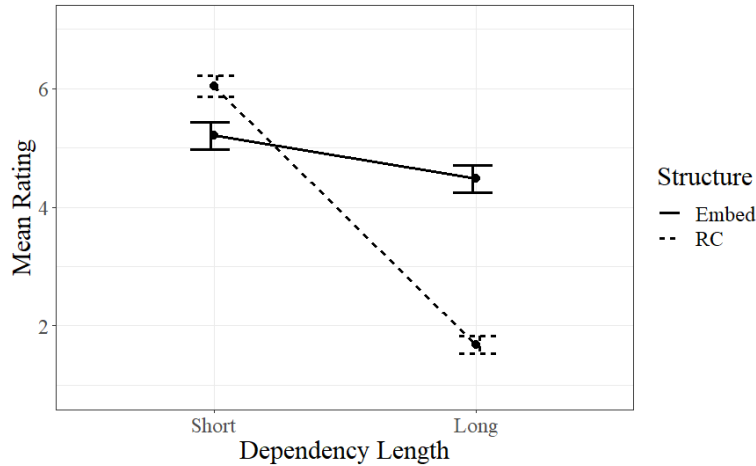

Figure 5. Interaction plot for adjunct *wh*-questions without context

Finally, Figure 6 shows the results of the acceptability judgment task on adjunct *wh*-questions with a given context providing a set of focus alternatives. There is a significant Structure effect (Est = -2.77, se = 0.37, p < 0.001) and a significant Length effect (Est = 1.56, se = 0.44, p < 0.01). The Structure x Length interaction is also significant (χ^2 = 27.73, p < 0.001). The superadditivity of adjunct *wh*-questions is present with or without the prior context.
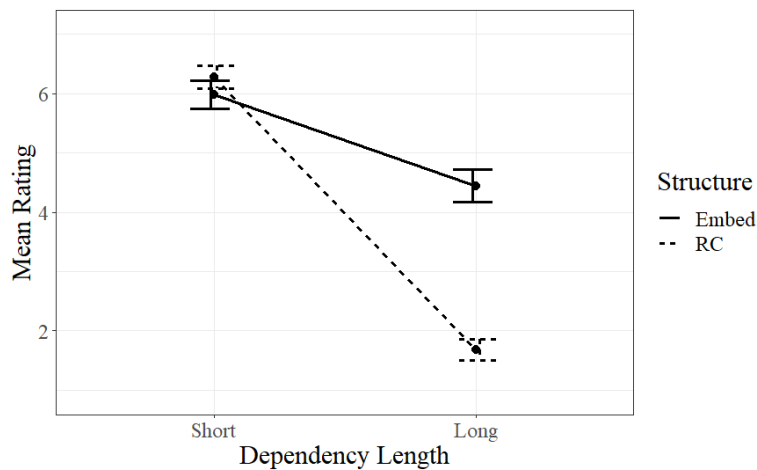

Figure 6. Interaction plot for adjunct *wh*-questions with context

To further test the effect of Context on argument and adjunct questions, two linear mixed-effects models were constructed for each Wh-category condition. The model structure is (Rating ~ Context*Structure*Length + (1|Item) + (1|Participant)). The three-way interaction of Context x Structure x Length is not significant for either the argument *wh*-questions (χ^2 = 0.52, p > 0.2) or

579

the adjunct *wh*-questions ($\chi^2 = 0.96$, p > 0.2). This result is expected for the adjunct *wh*-questions, as the superadditivity interaction on the target sentence is strong with or without context (see Figures 5 and 6). The lack of a reliable three-way interaction for the argument *wh*-questions should be treated with some caution. It is possible that the current experiment does not have sufficient statistical power. The experiment takes a 2x2x2x2 design with 16 conditions in total. The Context variable is between-subject, while the other variables are within-subjects. Therefore, each participant is exposed to eight conditions and, on average, three trials for each condition. A high-powered replication of the current study is needed to draw stronger conclusions.

We also examined the argument-adjunct asymmetry under the Context and No-context conditions separately. For the No Context condition, there is a significant three-way interaction of Wh-category x Structure x Length ($\chi^2 = 23.58$, p < 0.001). This three-way interaction shows that the superadditivity is greater for adjunct *wh*-questions than for argument *wh*-questions, which aligns with our observations from Figures 3 and 5. The adjunct *wh*-questions are more sensitive to RC-islands than argument *wh*-questions, even without context. The same argument-adjunct asymmetry is found for the With Context condition (three-way interaction: $\chi^2 = 4.74$, p < 0.05).

5. Discussion. Although the current study does not have sufficient power for strong statistical conclusions, preliminary results have some implications on two research topics – argument-adjunct asymmetry in Mandarin and the discourse account of island effects. The results of this experiment support a weak version of argument-adjunct asymmetry. Specifically, argument *wh*-questions in Chinese are still sensitive to RC-islands, contra previous studies (Huang 1982; Tsai 1994, 1999), but the island effect of argument *wh*-questions is more context-sensitive than adjuncts. When no context is provided, there is a weak but significant superadditivity in argument *wh*-questions and there is a strong superadditivity in adjunct *wh*-questions; however, when there is supportive context, the island effect disappears for argument but not for adjunct *wh*-questions. The stimuli of our no-context conditions were largely adopted from Lu et al. (2020). Although we replicate their basic finding that, for out-of-the-blue sentences, both argument and adjunct *wh*-questions demonstrate a superadditivity effect, the effect on argument *wh*-questions was much stronger in their study. One possible explanation is that our stimuli are syntactically less complex. The stimuli used in Lu et al. (2020) contain one more level of embedding than ours (see Tables 1 and 3). It is possible that with the additional level of embedding, it was even more difficult for participants to construct a focus-background distinction.

Our results provide some support for discourse indeed playing a role in modulating island sensitivity. Our hypothesis primarily focused on how the availability of alternatives can improve the acceptability of argument *wh*-questions with islands, at least in Mandarin. It is important to note, however, that the supportive context only improved the acceptability of argument *wh*-questions, not adjunct ones. We sketch two possibilities that may explain the robust sensitivity of adjunct *wh*-questions to islands.

Firstly, it is likely that the sources of the island effect are heterogeneous, including syntax, discourse, and processing factors. The island effect on adjunct *wh*-questions could be primarily driven by syntax with a relatively minimal role of discourse effect, which plays a larger role in argument *wh*-questions. The variations in different types of questions and different kinds of island structures create a complex problem of islands that not one approach can completely explain on its own. The contrast we observed in this study between argument and adjunct *wh*-questions could potentially lead to deeper insight into a more precise delineation of how different factors contribute to the island effect.

There is also a second possible explanation for the null effect of context in adjunct *wh*-questions. Argument and adjunct *wh*-questions require different types of alternatives. *Wh*-arguments have sets of alternatives that are clear entities. For instance, in a "who"-question, the alternatives contain individuals, such as {the manager, the cashier}. In a "what"-question, the alternatives contain entities, such as {elephants, rhinos}. These alternatives are relatively simple and they correspond to concepts people have experience with, and therefore stronger memory traces in long-term memory. On the other hand, the alternatives for "why"-questions are more complex. For an event that "the boy bought comic books", the possible reasons for this event might be a set like {because he is celebrating his birthday, because he received allowances from his parents, etc.}. These are alternatives at the proposition level, and the representation of these alternatives in our memory would be more complex and consequently demand more resources to maintain or/and access. The complexity of the set of alternatives may hinder the readers' access to it, even when the alternatives are explicitly presented in the context, reducing the effectiveness of context on adjunct *wh*-questions. This difference resonates with the discussion by Reinhart (1998), who noted that *wh*-arguments, but not *wh*-adjuncts, denote sets of individual entities and can be interpreted by choice functions.

A potential drawback of this experiment is its task complexity and the lack of a comprehension question. In the Context condition, in order to introduce the set of alternatives, the context was made quite complex. For each trial in the Context condition, participants need to read the context, maintain it in memory, and judge the acceptability of a target sentence. This process requires the participant to keep track of a lot of linguistic materials. There may also be individual differences in how engaged participants were during the experiment. All of these may have potentially impacted the effectiveness of context on the target sentences.

6. Conclusion. This study aims to investigate the role of discourse context in modulating the island effect of *wh*-questions in Mandarin Chinese. The interpretation of a *wh*-question requires the construction of a set of alternatives. We propose that the difficulty of extracting a *wh*-element out of an island is at least partially explained by the availability of the set of relevant alternatives people can construct. The acceptability rating results from the current study suggest that although argument *wh*-questions are subject to this discourse constraint, adjunct ones are not. Our findings are to some extent in line with the traditionally assumed argument-adjunct asymmetry; but in the meantime, crucially different from previous analysis, our results also suggest that the root cause of the argument-adjunct asymmetry involves a complex interplay between syntax and discourse constraints, instead of syntactic constraints alone.

References

Abeillé, Anne, Barbara Hemforth, Elodie Winckel & Edward Gibson. 2020. Extraction from subjects: Differences in acceptability depend on the discourse function of the construction. *Cognition* 204. 104293. https://doi.org/10.1016/j.cognition.2020.104293.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–85. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1977. On WH-movement. In Peter W. Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 71– 132. New York: Academic Press.

Erteschik-Shir, Nomi. 1973. *On the nature of island constraints.* Cambridge, MA: MIT dissertation.

Erteschik-Shir, Nomi. 1986. *Wh*-questions and focus. *Linguistics and Philosophy* 9(2). 117–149. https://www.jstor.org/stable/25001238.

Goldberg, Adele E. 2013. Backgrounded constituents cannot be "extracted". In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and island effects*, 221–238. Cambridge, UK: Cambridge University Press.

Huang, Chin-Te James. 1982. *Logical relations in Chinese and the theory of grammar.* Cambridge, MA: MIT dissertation.

Keshev, Maayan & Aya Meltzer-Asscher. 2019. A processing-based account of subliminal wh-island effects. *Natural Language & Linguistic Theory* 37. 621–657. https://doi.org/10.1007/s11049-018-9416-1.

Kluender, Robert & Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8(4). 573–633. https://doi.org/10.1080/01690969308407588.

Kuno, S. 1987. *Functional syntax: Anaphora, discourse and empathy*. Chicago, IL: University of Chicago Press.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26. https://doi.org/10.18637/jss.v082.i13.

Liu, Yingtong, Elodie Winckel, Anne Abeillé, Barbara Hemforth & Edward Gibson. 2022. Structural, functional, and processing perspectives on linguistic island effects. *Annual Review of Linguistics* 8. 495–525. https://doi.org/10.1146/annurev-linguistics-011619-030319.

Lu, Jiayi, Cynthia K. Thompson & Masaya Yoshida. 2020. Chinese *wh*-in-situ and islands: A formal judgment study. *Linguistic Inquiry* 51(3). 611–623. https://doi.org/10.1162/ling_a_00343.

Reinhart, Tanya. 1998. *Wh*-in-situ in the framework of the minimalist program. *Natural Language Semantic* 6(1). 29–56. https://doi.org/10.1023/A:1008240014550.

Rizzi, Luigi. 1982. *Issues in Italian syntax.* Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110883718.

Rooth, Mats. 1985. *Association with focus.* Amherst, MA: University of Massachusetts, Amherst dissertation.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116. https://doi.org/10.1007/BF02342617/.

Ross, John R. 1967. *Constraints on variables in syntax.* Cambridge, MA: MIT dissertation.

Sprouse, Jon, Matt Wagers & Colin Phillips. 2012. A test of the relation between working-memory capacity and syntactic island effects. *Language* 88(1). 82–123. https://www.jstor.org/stable/41348884.

Sprouse, Jon, Ivano Caponigro, Ciro Greco & Carlo Cecchetto. 2016. Experimental syntax and the variation of island effects in English and Italian. *Natural Language & Linguistic Theory* 34. 307–344. https://doi.org/10.1007/s11049-015-9286-8.

Tian, Qilin, Myung-Kwan Park & Xiaodong Yang. 2022. Mandarin Chinese *wh*-in-situ argument–adjunct asymmetry in island sensitivity: Evidence from a formal judgment study. *Frontiers in Psychology* 13. 954175. https://doi.org/10.3389/fpsyg.2022.954175.

Tsai, Wei-Tien Dylan. 1994. On nominal islands and LF extraction in Chinese. *Natural Language & Linguistic Theory* 12(1). 121–175. https://doi.org/10.1007/BF00992747.

Tsai, Wei-Tien Dylan. 1999. On lexical courtesy. *Journal of East Asian Linguistics* 8(1). 39–73. https://www.jstor.org/stable/20100754.

Xiang, Ming, Brian Dillon, Matt Wagers, Fengqin Liu & Taomei Guo. 2014. Processing covert dependencies: An SAT study on Mandarin *wh*-in-situ questions. *Journal of East Asian Linguistics* 23. 207–232. https://doi.org/10.1007/s10831-013-9115-1.

Xiang, Ming, Suiping Wang & Yanling Cui. 2015. Constructing covert dependencies – The case of Mandarin *wh*-in-situ dependency. *Journal of Memory and Language* 84. 139–166. https://doi.org/10.1016/j.jml.2015.05.006.

Zehr, Jeremy & Florian Schwarz. 2018. PennController for internet based experiments (IBEX). https://doi.org/10.17605/OSF.IO/MD832.